Journal of Multivariate Analysis **77**, 286–294 (2001)
doi:10.1006/jmva.2000.1943, available online at http://www.idealibrary.com on IDEAL®

Certain Characterizations of Normal Distribution via Transformations

G. G. Hamedani

Marquette University

and

Hans Volkmer

University of Wisconsin–Milwaukee

Received May 6, 1999; published online April 6, 2001

orthogonal transformation. Less well known are nonlinear transformations with the above-mentioned property. In this work we present nonlinear transformations preserving normality, which are more general than the existing ones in the literature. © 2001 Academic Press

AMS 1991 subject classification: 62E10.

Key words and phrases: normal characterizations; ergodic transformations; measure preserving transformations.

1. INTRODUCTION

Yu. V. Linnik presented three nonlinear transformations, (1), (2), and (3) given below, which transform a normal sample into itself. We state his results here, as they appeared in [1], for the sake of completeness.

THEOREM L_1. *Let X_1 and X_2 be two independently and normally distributed random variables with mean* 0 *and variance σ^2. Define two random variables Y_1 and Y_2 by the relation,*

$$Y_1 + iY_2 = (X_1 + iX_2)^k \, (X_1^2 + X_2^2)^{-(k-1)/2}. \tag{1}$$

Here k is a positive integer and $i = \sqrt{-1}$.

Then Y_1 and Y_2 are also independently and normally distributed with mean 0 *and variance σ^2.*

For $k = 2$ we have

$$Y_1 = (X_1^2 - X_2^2)(X_1^2 + X_2^2)^{-1/2}, \qquad Y_2 = 2X_1 X_2 (X_1^2 + X_2^2)^{-1/2},$$

and for k odd, (1) *gives rational transformations.*

0047-259X/01 $35.00

THEOREM L_2. *Let $X_1, X_2, ..., X_n$ be a sample from a population whose distribution function is normal with mean 0 and variance σ^2. Let B be an $n \times n$ matrix of the form*

$$B = \begin{vmatrix} M & 0 \\ 0 & I \end{vmatrix},$$

where M is an $\ell \times \ell$ orthogonal matrix whose elements a_{ij} are rational functions of the $n-\ell$ variables $X_{\ell+1}, X_{\ell+2}, ..., X_n$ $(2 \leqslant \ell \leqslant n-1)$, $a_{ij} = a_{ij}(X_{\ell+1}, X_{\ell+2}, ..., X_n)$, $i, j = 1, 2, ..., \ell$, while I is the $(n-\ell) \times (n-\ell)$ identity matrix. Let $(X_1, X_2, ..., X_n)^T$ be the column vector which corresponds to the elements of the sample and let

$$(Y_1, Y_2, ..., Y_n)^T = B(X_1, X_2, ..., X_n)^T. \tag{2}$$

Then the random variables $Y_1, Y_2, ..., Y_n$ are also independently and normally distributed with mean 0 and variance σ^2.

THEOREM L_3. *Let X_1 and X_2 be two independently and normally distributed random variables with mean 0 and variance σ^2. Define two random variables Y_1 and Y_2 by the relations*

$$\begin{aligned} Y_1 &= X_1 \cos a(X_1^2 + X_2^2) + X_2 \sin a(X_1^2 + X_2^2), \\ Y_2 &= -X_1 \sin a(X_1^2 + X_2^2) + X_2 \cos a(X_1^2 + X_2^2), \end{aligned} \tag{3}$$

where a is a constant. Then Y_1 and Y_2 are independently and normally distributed with mean 0 and variance σ^2.

In [2], it is stated that "nonlinear transformations preserving normality can be set up in the form of entire functions; for instance, the matrix B described above could have entire functions as elements. Another form of transformation is given by the formula

$$\begin{aligned} Y_1 &= X_1 \cos[\varphi(X_1^2 + X_2^2)] + X_2 \sin[\varphi(X_1^2 + X_2^2)], \\ Y_2 &= -X_1 \sin[\varphi(X_1^2 + X_2^2)] + X_2 \cos[\varphi(X_1^2 + X_2^2)], \end{aligned} \tag{4}$$

where φ is some entire function."

In [2], it is also stated that "compositions of the transformations indicated above ((1)–(4)) again lead to transformations preserving the normality of the sample. However, the general problem of describing all rational, all algebraic, or all entire transformations which preserve normality is apparently very difficult. Only some isolated facts are known in this area (cf. the note [3]). Among these we note a result due to Eidlin, communicated by him to the authors.

THEOREM E.[1] *Let $\sigma > 0$ be a given number. Consider a random sample $(X_1, X_2, ..., X_n)$ with every $X_j \sim N(0, \sigma^2)$. Every algebraic transformation preserving normality of such a sample also preserves spheres $x_1^2 + x_2^2 + \cdots + x_n^2 = r^2$ and also Lebesgue measure.*"

In Theorem L_i, $i = 1, 2, 3$, and Theorem E it is assumed that $E(X_i) = 0$. If $E(X_i) = \mu \neq 0$, then all the transformations given above require the obvious modifications. It is also clear that (3) is a special case of (4).

Beer and Lukacs [1] proved the converses of Theorem L_i, $i = 1, 2, 3$, and hence obtained in this way three characterizations of the normal distribution, which are stated below.

THEOREM BL_1. *Let X_1 and X_2 be i.i.d. random variables and suppose that their distribution function is absolutely continuous and has a continuous pdf (probability density function) $f(x)$. Let Y_1 and Y_2 be determined by (1) and suppose that Y_1 and Y_2 are also i.i.d. with pdf $f(x)$. Then $f(x) = (\sigma\sqrt{2\pi})^{-1} \exp(-x^2/2\sigma^2)$.*

THEOREM BL_2. *Let $X_1, X_2, ..., X_n$ be i.i.d. random variables with distribution function $F(x)$. Let ℓ be an integer, $2 \leqslant \ell \leqslant n-1$, and let M be a nontrivial orthogonal $\ell \times \ell$ matrix (a matrix is called trivial if each row contains only one nonzero element) whose elements a_{ij} are rational functions of the $n-\ell$ variables $x_{\ell+1}, x_{\ell+2}, ..., x_n$ $(i, j = 1, 2, ..., \ell)$. Let B be an $n \times n$ matrix of the form*

$$B = \begin{vmatrix} M & 0 \\ 0 & I \end{vmatrix},$$

where I is the $(n-\ell) \times (n-\ell)$ identity matrix. Let $(Y_1, Y_2, ..., Y_n)^T$ be given by (2). Assume that the random variables $Y_1, Y_2, ..., Y_n$ are i.i.d. with distribution function $F(x)$. Then $F(x)$ is a normal distribution with zero mean.

THEOREM BL_3. *Let X_1 and X_2 be two i.i.d random variables with absolutely continuous distribution and continuous pdf $f(x)$. Let Y_1 and Y_2 be given by (3) with $a \neq 0$. Assume further that Y_1 and Y_2 are also i.i.d. with pdf $f(x)$. Then $f(x) = (\sigma\sqrt{2\pi})^{-1} \exp(-x^2/2\sigma^2)$.*

In the present work we establish two theorems which generalize and unify Theorem L_i, $i = 1, 2, 3$, and Theorem BL_i, $i = 1, 3$, respectively. Beer

[1] The present authors contacted A. M. Kagan regarding a proof of this theorem. He informed us that V. L. Eidlin passed away before publishing his proof and no one possesses a proof of this theorem.

and Lukacs [1] did not provide a characterization of the normal distribution based on transformation (4). Our results will include (4) as a special case as well. It will be seen that transformations (1)–(4) preserving the normality of a sample are special cases of our transformation given below.

2. PRELIMINARY RESULTS

Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. A measurable transformation $T: \Omega \to \Omega$ preserves the measure μ if $\mu(T^{-1}A) = \mu(A)$ for all sets $A \in \mathscr{A}$. In this case μ is called an invariant measure for T. If T preserves the finite measure μ, then T is called ergodic on $(\Omega, \mathscr{A}, \mu)$ if the only sets A in $\mathscr{A}$ with $T^{-1}A = A$ satisfy $\mu(A) = 0$ or $\mu(A) = \mu(\Omega)$. These definitions can be found in [4].

The following lemma on the uniqueness of invariant measures will be used later.

LEMMA 1. *Let $T: \Omega \to \Omega$ be ergodic on the probability space $(\Omega, \mathscr{A}, \mu)$. Let $g: \Omega \to \mathbb{R}$ be an integrable function inducing the probability measure $\nu(A) = \int_A g\, d\mu$. If T preserves the measure ν, then $\mu = \nu$.*

Proof. Define $A = \{w \in \Omega \mid g(w) < 1\}$. Using the invariance property of ν, we obtain

$$\begin{aligned} \nu(A \cap T^{-1}A) + \nu(A \backslash T^{-1}A) &= \nu(A) = \nu(T^{-1}A) \\ &= \nu(A \cap T^{-1}A) + \nu(T^{-1}A \backslash A). \end{aligned}$$

Thus

$$\int_{A \backslash T^{-1}A} g\, d\mu = \int_{T^{-1}A \backslash A} g\, d\mu.$$

Since T preserves μ, we also have $\mu(A \backslash T^{-1}A) = \mu(T^{-1}A \backslash A)$. Since $g < 1$ on $A \backslash T^{-1}A$ and $g \geqslant 1$ on $T^{-1}A \backslash A$, we obtain $\mu(A \backslash T^{-1}A) = \mu(T^{-1}A \backslash A) = 0$ and so $\mu(T^{-1}A\, \Delta A) = 0$. Since T is ergodic on $(\Omega, \mathscr{A}, \mu)$, this implies that $\mu(A) = 0$ or $\mu(A) = 1$; see [4, Theorem 1.5 (ii)]. If $\mu(A) = 1$, then

$$1 = \nu(\Omega) = \int_A g\, d\mu < \mu(A) = 1,$$

a contradiction. Hence $\mu(A) = 0$. Similarly, we can show that $\mu(B) = 0$ for $B = \{w \in \Omega \mid g(w) > 1\}$. Thus $g(w) = 1$ for almost all w. This completes the proof.

Let $\|x\|$ denote the Euclidean norm of a vector x in $\mathbb{R}^n$. For $r>0$, let $S_r=\{x\in\mathbb{R}^n \mid \|x\|=r\}$. Let λ denote the Lebesgue measure on $\mathbb{R}^n$ and let λ_r denote the Lebesgue surface measure on S_r. Note that if $h:\mathbb{R}^n\to\mathbb{R}$ is integrable, then its restriction $h_r: S_r\to\mathbb{R}$ is integrable with respect to λ_r for almost all $r>0$ and

$$\int_{\mathbb{R}^n} h\,d\lambda=\int_0^\infty \int_{S_r} h_r\,d\lambda_r\,dr. \tag{5}$$

LEMMA 2. *Let $T:\mathbb{R}^n\to\mathbb{R}^n$ be a measurable transformation such that $\|Tx\|=\|x\|$ for all $x\in\mathbb{R}^n$. Let $h:\mathbb{R}^n\to\mathbb{R}$ be an integrable function. Then the following statements are equivalent*:

(i) *T preserves the measure ν defined by*

$$\nu(A)=\int_A h\,d\lambda;$$

(ii) *for almost all $r>0$, T_r preserves the measure ν_r on S_r defined by*

$$\nu_r(B)=\int_B h_r\,d\lambda_r.$$

Proof. Assume (ii) holds. Let A be a measurable subset of $\mathbb{R}^n$. For $r>0$, let $A_r=A\cap S_r$. Using (5) and (ii) we obtain

$$\nu(A)=\int_A h\,d\lambda=\int_0^\infty \nu_r(A_r)\,dr=\int_0^\infty \nu_r(T^{-1}A_r)\,dr=\nu(T^{-1}A),$$

since $T^{-1}(A)\cap S_r=T^{-1}(A_r)$. So (i) is proved.

Assume (i) holds. We choose a countable semialgebra Φ of measurable subsets of S_1 which generates the σ-algebra of measurable sets in S_1. Let $B\in\Phi$, and for $R>0$ define $A_R=\{x\in\mathbb{R}^n \mid \frac{x}{\|x\|}\in B,\ \|x\|\leqslant R\}$. For $r>0$, define $B_r=\{x\in\mathbb{R}^n \mid \frac{x}{\|x\|}\in B,\ \|x\|=r\}$. Then using (5) and (i) we obtain

$$\int_0^R \nu_r(B_r)\,dr=\nu(A_R)=\nu(T^{-1}A_R)=\int_0^R \nu_r(T^{-1}B_r)\,dr.$$

Since this is true for all $R>0$, we find that

$$\nu_r(B_r)=\nu_r(T^{-1}B_r)$$

for almost all $r>0$. Since Φ is a countable collection, this is true for all $B\in\Phi$ for almost all $r>0$. By [4, Theorem 1.1], T_r preserves the measure ν_r for almost all $r>0$, completing the proof.

Finally, we say that a measurable function $h : \mathbb{R}^n \to \mathbb{R}$ is radially symmetric if there is a function $k : [0, \infty) \to \mathbb{R}$ such that $h(x) = k(\|x\|)$ for almost all $x \in \mathbb{R}^n$.

THEOREM 1. *Assume that the measurable transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ has the following two properties*:

(a) $\|Tx\| = \|x\|$ *for all* $x \in \mathbb{R}^n$;

(b) *for almost all $r > 0$, the transformation $T_r : S_r \to S_r$ preserves the Lebesgue surface measure λ_r on S_r.*

Then T preserves every measure

$$\nu(A) = \int_A h \, d\lambda$$

defined by a radially symmetric integrable function $h : \mathbb{R}^n \to \mathbb{R}$.

Proof. It follows immediately from Lemma 2 because the measures ν_r and λ_r are equal modulus a constant factor (depending on r).

The following theorem is a converse of Theorem 1.

THEOREM 2. *Assume that the measurable transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ has properties* (a) *and* (b) *of Theorem* 1 *as well as the following property*:

(c) *for almost all $r > 0$, T_r is ergodic with respect to λ_r.*

Further, assume that T preserves the measure $\nu(A) = \int_A h \, d\lambda$ defined by an integrable function $h : \mathbb{R}^n \to \mathbb{R}$. Then h is radially symmetric.

Proof. By Lemma 2, T_r preserves measure ν_r for almost all $r > 0$. By our assumptions, T_r also preserves the Lebesgue surface measure λ_r and is ergodic. Now Lemma 1 implies that $\nu_r = k(r) \, \lambda_r$ for almost all $r > 0$ where $k : [0, \infty) \to \mathbb{R}$ is defined by $k(r) = \nu_r(S_r)/\lambda_r(S_r)$. But this implies that $h(x) = k(\|x\|)$ for almost all $x \in \mathbb{R}^n$. So h is radially symmetric.

3. FINAL RESULTS

Now we employ Theorems 1 and 2 of the preceeding section to establish our main results.

THEOREM 3. *Let $X_1, X_2, ..., X_n$ be i.i.d. standard normal random variables. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a transformation with properties* (a) *and* (b).

Then the random variables $Y_1, Y_2, ..., Y_n$ *defined by* $(Y_1, Y_2, ..., Y_n) = T(X_1, X_2, ..., X_n)$ *are also i.i.d.* $N(0, 1)$.

Proof. The joint pdf of $X_1, X_2, ..., X_n$ is

$$f(x_1, x_2, ..., x_n) = (2\pi)^{-n/2} \exp\left(-\frac{\|x\|^2}{2}\right).$$

Let ν be the probability measure induced by f. By Theorem 1, T preserves the measure ν. But this just means that the joint pdf of $Y_1, Y_2, ..., Y_n$ is f.

THEOREM 4. *Let* $X_1, X_2, ..., X_n$ *be i.i.d. with continuous pdf* $f(x)$. *Let* $T: \mathbb{R}^n \to \mathbb{R}^n$ *be a transformation with properties* (a), (b), *and* (c). *Let the random variables* $Y_1, Y_2, ..., Y_n$ *be defined by* $(Y_1, Y_2, ..., Y_n) = T(X_1, X_2, ..., X_n)$. *If* $Y_1, Y_2, ..., Y_n$ *are also i.i.d. with pdf* $f(x)$, *then there is* $\sigma > 0$ *such that* $f(x) = (\sigma\sqrt{2\pi})^{-1} \exp(-x^2/2\sigma^2)$.

Proof. Define $g(x_1, x_2, ..., x_n) = f(x_1)\, f(x_2) \cdots f(x_n)$. By our assumption, T preserves the measure induced by g. By Theorem 2, g is radially symmetric. Since f is continuous, this implies easily that $f(x) = a \exp(bx^2)$ where a and b are constants; see [1].

Remarks. (a_1) Theorem 3 includes Theorem L_i, $i = 1, 2, 3$, as special cases. In Theorem L_1, we have $n = 2$ and the transformation $(y_1, y_2) = T(x_1, x_2)$ is given by

$$y_1 + iy_2 = (x_1 + ix_2)^k\, (x_1^2 + x_2^2)^{-(k-1)/2},$$

where k is a positive integer. Clearly T satisfies the property (a). The induced transformation $T_r : S_r \to S_r$ is given by $\phi \to k\phi$ in polar coordinates. It is easy to see that T_r preserves the arclength measure on S_r.

In Theorem L_2, we set $n = \ell + m$ with positive integers ℓ and m. The transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is given by $(y_1, y_2, ..., y_n) = T(x_1, x_2, ..., x_n)$, where

$$(y_1, y_2, ..., y_\ell)^T = M(x_{\ell+1}, x_{\ell+2}, ..., x_n)(x_1, x_2, ..., x_\ell)^T$$

and

$$y_i = x_i \qquad \text{if} \quad i = \ell+1, \ell+2, ..., n.$$

The $\ell \times \ell$ matrix M is orthogonal with entries that may depend on $x_{\ell+1}, x_{\ell+2}, ..., x_n$. Clearly T satisfies property (a). We also see that T is one-to-one and onto. For fixed $r > 0$, let B be a measurable subset of S_r. Let $A = T^{-1}(B)$ so that $B = T(A)$. For fixed $(x_{\ell+1}, x_{\ell+2}, ..., x_n)$ with norm s less than r, let A' be the set of all $(x_1, x_2, ..., x_\ell)$ such that $(x_1, x_2, ..., x_n)$

lies in A. Then A' is a measurable subset of $S_u^{\ell-1}$ where $u=\sqrt{r^2-s^2}$. Define B' similarly. By the definition of T and the orthogonality of M, the $\ell-1$ dimensional surface measures of A' and B' agree. It follows easily that the $n-1$ dimensional surface measures of A and B agree. This construction can best be illustrated by taking $n=3$ and $\ell=2$. Then T restricted to S_r describes "rotations" of the sphere about the x_3-axis where the rotation angle on a given surface (given by $x_3=\text{constant}$) may depend on x_3.

In Theorem L_3, $n=2$ and $(y_1, y_2)=T(x_1, x_2)$ is defined by

$$y_1=x_1\cos ar^2+x_2\sin ar^2, \qquad y_2=-x_1\sin ar^2+x_2\cos ar^2,$$

where a is a constant and $r=\|x\|$. Clearly T satisfies the property (a). For fixed $r>0$, the transformation $T: S_r\to S_r$ is a rotation by the angle ar^2. Thus T also satisfies property (b).

(a$_2$) Theorem BL_1 follows from Theorem 4 because the map $T_r: S_r \to S_r$ given by $\phi\to k\phi$ is ergodic for $r>0$ and every integer $k\geqslant 2$; see [4, Corollary 1.10.1]. Theorem BL_3 is also a special case of Theorem 4 because the rotation T_r given by $\phi\to\phi+ar^2$ is ergodic if and only if ar^2 is not a rational multiple of 2π; see [4, Theorem 1.8]. Thus, for almost all $r>0$, T_r is ergodic.

Theorem BL_2 cannot be obtained directly from Theorem 4 as the following example demonstrates. Let T be a rotation about the z-axis by the angle $2\pi/3$ in R^3. Then with $\ell=2$, $n=3$, Theorem BL_2 applies because T is nontrivial. Theorem 4 does not apply because the rotations T_r are not ergodic for any $r>0$. The transformation T_r is ergodic if and only if every measurable function h defined on S_r and satisfying $hT_r=h$ is constant a.e. [4, Theorem 1.6]. In order to prove Theorem 4, we need this property only for functions h which are of the form $h(x_1, ..., x_n)=g(x_1)\cdots g(x_n)$. If we weaken assumption (c) of Theorem 4 in this way, we can also obtain Theorem BL_2 from Theorem 4.

(a$_3$) The transformation (4) clearly has properties (a) and (b) whenever $\varphi: [0,\infty)\to\mathbf{R}$ is measurable. So a theorem analogous to Theorem L_3 holds for this transformation. The transformation T has the property (c) if we require that the set of all $t\geqslant 0$ for which $\varphi(t)$ is a rational multiple of 2π is a null set. Under this condition, a theorem analogous to Theorem BL_3 holds.

(a$_4$) Combining (1) and (4), we can consider the transformation $T:\mathbb{R}^2\to\mathbb{R}^2$ which keeps the radius r fixed and changes the polar angle θ to $n(r)\,\theta+a(r)$, where $n(r)$ is a measurable function on $(0,\infty)$ with values in $\mathbb{N}^+$ and $a(r)$ is a real measurable function defined on $(0,\infty)$. This transformation satisfies properties (a) and (b) of Theorem 1.

ACKNOWLEDGMENT

The authors thank Eric Key for helpful discussions. We are also grateful to two anonymous referees for their valuable comments.

REFERENCES

1. S. Beer and E. Lukacs, Characterization of the normal distributions, *J. Appl. Probab.* **10** (1973), 100–108.
2. A. M. Kagan, Yu. V. Linnik, and C. R. Rao, “Characterization Problems in Mathematical Statistics,” Wiley, New York, 1973.
3. Yu. V. Linnik and V. L. Eidlin, On analytical transformations of normal vectors, *Teoriia Veroiatn. Prim.* **13** (1968), 751–754.
4. P. Walters, “An Introduction to Ergodic Theory,” Springer-Verlag, New York, 1982.